Library & Information Science Research 36 (2014) 211–224

Contents lists available at ScienceDirect

Library & Information Science Research

Benchmarking local public libraries using non-parametric frontier methods: A case study of Flanders

Jesse Stroobants *, Geert Bouckaert

Public Governance Institute, University of Leuven, Faculty of Social Sciences, Parkstraat 45 bus 3609, B-3000 Leuven, Belgium

ARTICLE INFO

Available online 28 October 2014

Keywords:
Benchmarking
Performance information
Efficiency
Free disposal hull (FDH)
Data envelopment analysis (DEA)

ABSTRACT

Being faced with significant budget cuts and continual pressure to do more with less, issues of efficiency and effectiveness became a priority for local governments in most countries. In this context, benchmarking is widely acknowledged as a powerful tool for local performance management and for improving the efficiency and effectiveness of local service delivery. Performance benchmarking exercises are regularly carried out using ratio analysis, by comparing single indicators. Since this approach offers only limited assessments in absolute terms, it is difficult for decision-makers to track and improve overall performance. Therefore, the use of non-parametric frontier methods, namely free disposal hull (FDH) and data envelopment analysis (DEA) is presented as an alternative technique for benchmarking the performance of organizations in relative terms. The potential applications and strengths of these non-parametric frontier methods for benchmarking the efficiency of local public services are highlighted by applying FDH and DEA techniques to the local public libraries in Flanders. Incorporating all possible paths of expansion – both in space and in time – enables a focus on sustainability within efficiency benchmarking.

1. Introduction

The current economic and financial situation puts public sector performance under pressure, both in Flanders and abroad. In fact, the budgetary space for the different Belgian governments is shrinking, while the demand for qualitative public services keeps rising (Troupin, Stroobants, & Steen, forthcoming). Also local governments in Flanders cannot avoid the consequences of the economic crisis and decreasing revenues (Belfius, 2013). In its policy statement on Internal Administration 2009–2014, the Flemish Government indicated that over the coming years there is no policy space for creating important financial incentives, meaning that local governments themselves have to take responsibility for improving their efficiency and effectiveness in order to perform better, and to provide more and better services, with less means (Bourgeois, 2009).

One of today's main instruments for measuring and evaluating performance, as a tool for identifying and adopting more efficient and effective practices, is *benchmarking* (Fenna, 2012). Benchmarking involves placing an entity's performance in context by comparing performance with standards, with figures for the same measures in previous reporting periods, or with performance results achieved by others (Ammons, 2012). Applied in the public sector, benchmarking is defined as contextualising the current performance of a public sector organization by comparing it with other (similar) organizations or its own past, with an intent to improve (Askim, Johnsen, & Christophersen, 2008; Berg, 2010).

Benchmarking methods for performance comparisons are mostly developed and introduced by practitioners. Many practitioners use simple techniques rather than analytical methods (Talluri, 2000). Besides these more simple benchmarking approaches, non-parametric frontier techniques can be effective and alternative methods for performance analysis and benchmarking when the measurement issue is considered in terms of (technical) efficiency.[1] For benchmarking local government performance, however, there is still a need to demonstrate the application and value of these non-parametric frontier methods: free disposal hull (FDH) and data envelopment analysis (DEA).

2. Problem statement

Benchmarking and performance comparisons of public sector organizations are usually conducted by using (a set of) indicators or performance measures, especially when carried out by practitioners and decision-makers in the policy arena (van Helden & Reichard, 2013).

* Corresponding author.
E-mail address: jesse.stroobants@soc.kuleuven.be (J. Stroobants).

[1] Technical efficiency is the ability to convert a certain bundle of inputs to the maximum possible amount of output (with current technology, as evidenced by the best performance observed) – or, alternatively, the situation in which as little input as possible is used in producing a certain amount of output (Agasisti & Johnes, 2009).

http://dx.doi.org/10.1016/j.lisr.2014.06.002
0740-8188/

Table 1
Efficiency analyses of libraries[a], specification of inputs and outputs.

Reference	Sample/technique	Inputs	Outputs
Kwack (1993)	20 university libraries/DEA	Library staff Area of library space Number of library books	Reader visits Book circulation
Chen (1997)	23 university and college libraries/DEA	Library staff Book acquisition expenditure Book collection Area of library space Seating Capacity	Reader visits Book circulation Reference and on-line research Annual service hours Reader satisfaction Interlending service
Mann (1997)	108 university libraries/DEA	Staff Total expenditures	Total volumes Total volumes added Current serials[b]
Vitaliano (1998)	184 public libraries/DEA	Total holdings of all items (books, audiovisual, maps, etc.) Total hours of operation per week New books purchased Total serial subscriptions currently active[b]	Annual total circulation of all library materials Number of reference questions answered
Sharma, Leung, and Zane (1999)	47 public libraries/DEA	Book collection Library staff Days open (in 8-hour days) Operating expenses	Book circulation Reader visits Reference transaction
Worthington (1999)	168 local public libraries/DEA	Gross library expenditure	Number of library issues
Shim (2000)	95 academic libraries/DEA	Professional staff Support staff Student staff Volumes held Net volumes added Monographs purchased Total serials[b]	Total circulation Reference transactions Interlibrary lending Interlibrary borrowing Library instruction
Hammond (2002)	99 public library systems/DEA	Total opening hours per week Number of books and audio-visual material Acquisitions of new material Number of serial subscriptions[b]	Total number of items issued to borrowers over the year Number of enquiries processed Number of requests processed
Reichmann (2004)	118 university libraries/DEA	Staff (number of FTEs) Book materials held	Number of book materials added Weekly opening hours Circulation Serial subscription[b]
Jo, Park, Lee, and Yoon (2009)	26 university libraries/DEA	Number of employees Size of library Budget Number of books	Number of visitors Number of loan books Number of books for lending and borrowing
Miidla and Kikas (2009)	20 central public libraries/DEA	Yearly acquisition expenditures Yearly salary expenditures Collection size Floor area	Number of readers Number of loans
Reichmann and Sommersguter-Reichmann (2010)	68 university libraries/DEA	Number of employees (FTEs) Total number of book materials	Number of book materials added Total Circulation Number of serial subscriptions[b]
Noh (2011)	89 university libraries/DEA	Budget Number of librarians Number of books Number of serials Number of e-journals Web databases Number of e-books Number of computers Internally developed database units	Number of circulation books Number of users Number of website visits Number of database users
De Witte and Geys (2011)	290 local public libraries/FDH & DEA	Personnel expenditures Operating expenditures Infrastructure	Opening hours per week Youth books Fiction and non-fiction books Media (CD, DVD, VHS, CD-ROM)
De Witte and Geys (2012)	291 local public libraries/FDH	Personnel expenditures Operating expenditures Infrastructure	Youth books Fiction and non-fiction books Media (CD, DVD, VHS, CD-ROM) Young borrowers (<16 years) Total book circulation Media circulation
Simon, Simon, and Arias (2011)	34 university libraries/DEA	Personnel Total surface area Total expenditures on bibliographic-related materials (monographs, serial subscriptions, access to electronic resources, etc.)	Total circulation Inter-library loans Number of documents downloaded Number of monographs Number of serial subscriptions Number of seats Service hours

Table 1 *(continued)*

Reference	Sample/technique	Inputs	Outputs
de Carvalho, Jorge, Jorge, Russo, and de Sá (2012)	37 university libraries/DEA	Number of employees Area Number of volumes	Consultations Loans Enrolments User traffic

[a]Public and academic; [b]Serials are all paper periodicals (i.e., journals, magazines, annual reports, etc.).

Performance indicators have much to recommend them; they focus on specific aspects of performance (e.g., efficiency or effectiveness), are readily measured and validated, and are easy to interpret (in isolation, at least). Indicators might therefore be useful from a managerial perspective (Smith & Street, 2005); and, expressed in absolute terms, also provide a good starting point for benchmarking organizational performance in a simple manner (both to track an entity's own performance over time and to compare performance against other similar entities or against a relevant standard) (Berg, 2010).

Despite their merits, however, there are some drawbacks to using performance indicators. First, they provide only an indirect or partial indication of performance. For instance, with efficiency as performance measure, indicators will be single-input, single-output (Woodbury, Dollery, & Rao, 2003). Second, they may provide conflicting results: an organization that appears to do well on one indicator may perform less successfully when considered on another (Smith & Street, 2005).

Besides benchmarking by means of indicators, frontier methods can be identified as alternative techniques for measuring and evaluating performance of a group of comparable entities (Cummins & Weiss, 2013). Different from the single-factor measures that reflect only partial aspects of performance, frontier techniques can be applied to assess overall performance by handling multiple inputs and outputs at the same time (Worthington & Dollery, 2000). Moreover, non-parametric frontier methods are proven to be a useful tool for assessing the relative efficiency of entities.

Although there is recent and abundant literature on non-parametric efficiency measurement, especially using DEA, few studies incorporate all possible paths of expansion – both in space and time. To address this apparent gap in knowledge, the application of the two non-parametric methods – both FDH and DEA – for benchmarking local government efficiency is demonstrated in this study using four case studies concerning public libraries as an area of local public service delivery. These case studies represent the different directions in which such efficiency analyses can be further developed: the number of inputs and/or outputs, or the number of entities and the time scale.

This holistic approach is expected to allow identification of sustainable, stable, and efficient entities on the one hand and efficiency improvement trajectories for inefficient entities on the other. Such knowledge, methodologically ordered and organized, can link the benchmarking phase to the subsequent step of *benchlearning*.

3. Literature review

As noted earlier, benchmarking in the public sector is primarily carried out by using sets of performance indicators. However, today – driven mainly by academic research – there is a growing tendency to benchmark the relative performance of public entities using frontier methods. This is also the case for evaluating local government performance, as listed by Kalb, Geys, and Heinemann (2012). Such exercises focus mainly on efficiency of public service delivery, rather than on effectiveness or quality. In any case, by including multiple inputs and/or outputs, these studies have attempted to analyse the overall performance of public sector organizations.

Recently, researchers have also applied non-parametric frontier methods to libraries, whether public or academic. An investigation of the library performance management literature shows that there are several studies regarding the use of FDH or DEA to measure and assess the efficiency of libraries. The existing literature is summarized in Table 1, indicating the sample of libraries studied, the frontier technique used and the multiple inputs and outputs selected to assess the more aggregate library performance.

From a methodological point of view, many authors fell back on DEA. Only De Witte and Geys (2011, 2012) had previously applied FDH for analyzing the efficiency of public libraries. This overview also shows that there exists a large variety of possible input and output indicators for measuring library efficiency. The most commonly used inputs are *personnel* (expressed in numbers or expenditures), *operational expenditures*, and *library collection* (number of items). The most selected output variable, by far, is *circulation* (number of loans). As the determination of inputs and outputs is an important step in efficiency evaluation, this literature review will also help by identifying suitable input and output indicators from the case studies discussed.

4. Methods

Non-parametric methods use linear programming to construct a piecewise frontier that envelops all observations (DMUs[2]) of the sample used, against which each DMUs efficiency can be evaluated. In this manner, the performance of entities is not measured in absolute terms (as is the case with using indicators), but assessed relatively against each other, in what can be considered a pure mode of benchmarking. Generally, two non-parametric frontier methods can be distinguished: FDH analysis and DEA.

4.1. Free disposal hull (FDH)

The FDH method imposes the least amount of restrictions on the data, as it only assumes free disposability of resources (an entity can use more inputs than technically necessary to produce a certain level of output) and strong disposability of outputs (an entity can generate less output than technically possible with a certain amount of resources). In other words, the disposability assumptions imply that an increase in inputs never results in a decrease in outputs, and that any reduction in outputs remains producible with the same amount of inputs (De Borger & Kerstens, 1996).

FDH determines efficiency scores for each DMU by comparing *best practices* in a set of benchmarked entities. More specifically, it analyses if there are observations that generate the same output with fewer resources (*input efficiency*) or more output with the same input (*output efficiency*). Comparisons are only made with existing observations in the set, unlike DEA where comparisons with virtual DMUs are also made (see below). The FDH frontier is a stairway-shaped curve connecting the efficient observations, as illustrated in Fig. 1 for a single-input (x), single-output (y) case. It can be noticed that, following the central hypothesis of the FDH approach, DMU e in Fig. 1 is found to be inefficient because both DMUs b and c generate more output with less input.

[2] Within the context of non-parametric frontier methods, entities or organisation units under study are called decision-making units (DMUs).

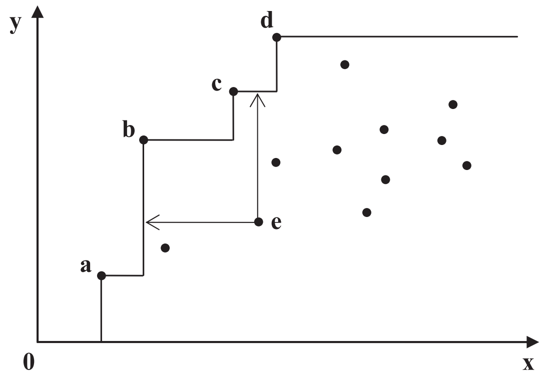


Fig. 1. Free disposal hull (FDH) efficiency frontier.

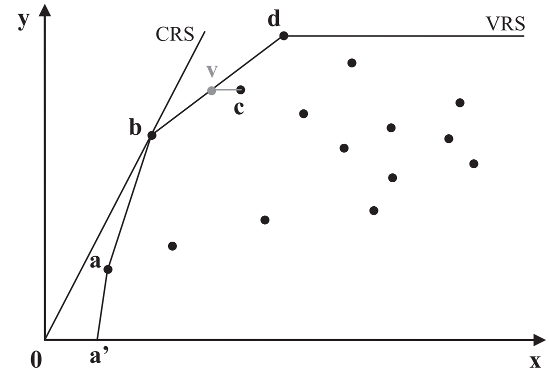


Fig. 2. Data envelopment analysis (DEA) efficiency frontier.

4.2. Data envelopment analysis (DEA)

DEA, the second non-parametric frontier approach, adds the assumption that linear combinations of the observed input-output bundles are also feasible. Hence, it assumes the existence of a convex production frontier, a hypothesis that is not required in the FDH approach. The term *envelopment* stems from the fact that the frontier envelops the set of observations. Fig. 2 illustrates the DEA frontier in a single-input (x) single-output (y) context. It may be noted that DMU c, considered as efficient by FDH, is now lying below the efficiency frontier. This is due to the convexity assumption, meaning that the efficiency of DMU c is not only ranked against the real performers (DMU b and d), but also evaluated against virtual units such as DMU v, being a linear combination of DMU b and d (Herrera & Pang, 2005). This makes the DEA approach more strict than FDH analysis, meaning that fewer DMUs will be found fully efficient and the efficiency scores will be lower (Afonso & St. Aubyn, 2005).

Within the DEA methodology, a further distinction is made between two efficiency frontiers based on the returns to scale assumption. Returns to scale indicate the increase in output produced from a proportional increase in all inputs. Increasing (respectively, decreasing) returns to scale indicate that an increase in the input resources generates a more (respectively, less) than proportionate increase in outputs. With constant returns, output increases by the same proportion as an increase of the inputs.

The original DEA models formulated by Farrell (1957) and Charnes, Cooper, and Rhodes (1978) assumed the existence of constant returns to scale (CRS), implying no scale effects in the size of operation. The DEA frontier under CRS is represented in Fig. 2 by a straight line extending from the origin through the efficient DMU (ray $0b$). By this standard, only DMU b would be rated efficient (Herrera & Pang, 2005). However, DEA models under the variable to scale (first increasing and then decreasing) assumption are more common today. The DEA frontier $a'abd$ reflects variable returns to scale (VRS), because only convex combinations of efficient DMUs form the best-practice frontier.[3]

5. Benchmarking by using FDH and DEA: Case studies of public libraries in Flanders

Four case studies were carried out to illustrate the application of non-parametric frontier techniques, FDH and DEA, for benchmarking local governments. These cases illustrate (a) what the possibilities are of non-parametric frontier methods for comparing performance between entities or within time; (b) which data requirements should be met; and (c) how the results have to be interpreted.

In the case studies, the frontier methods were applied on a specific domain of local public service delivery, namely public libraries in Flanders.[4] The choice for benchmarking public libraries was based on the existence of a large and detailed database, Bios2 (2013). This database contains more than 140 parameters on which the 309 public libraries in Flanders have reported since 2006, making it not only possible to compare performance over time for each public library, but also (and maybe more important in the context of local benchmarking) to compare between libraries for one or more performance parameters.

As non-parametric frontier techniques involve the use of linear programming methods to evaluate the relative performance of a set of DMUs, computer software is needed to apply these techniques. There are several packages to carry out standard FDH and DEA analyses (e.g., Banxia, DEAFrontier, DEA-Solver, Frontier Analyst). For these case studies, *DEAFrontier*[5] software, a user-friendly add-in for Microsoft Excel, was used.

5.1. Integrated benchmarking approach: Moving along three axes

Since non-parametric frontier methods would be used for benchmarking local public libraries in Flanders, the focus was consequently on efficiency as performance measure. Benchmarking the efficiency of public libraries can be done for a particular library in regular time intervals (i.e., comparing against its own performance over time), for a set of libraries (i.e., comparing against other entities), or for a certain selection of inputs and outputs as part of the efficiency measure.

To graphically illustrate how efficiency benchmarking can shift according the above-mentioned aspects, a three-dimensional space was developed (see Fig. 3) in which the three axes represent the possible benchmarking directions. This figure acts as the starting point, or base scheme, for the four benchmarking cases in this section.

On the X-axis, the number of inputs and outputs included in the benchmarking can vary, because an efficiency analysis with frontier methods is not limited to incorporating one single input (e.g., expenditures or staff) and one single output (e.g., number of loans), as is the case with an efficiency indicator. Instead, FDH and DEA can accommodate multiple inputs and multiple outputs simultaneously. Second, on the Y-axis, the number of DMUs in the benchmarking sample can differ. In principle, the efficiency of all 309 public libraries in Flanders could be compared relatively to each other. However, since it is a basic rule that only comparable entities should be benchmarked, it is recommended to compose a deliberate set of DMUs (e.g., the public libraries in the 13 central cities in Flanders).

[3] The acronyms CCR and BCC are used at times in reference to CRS and VRS models. The acronyms are formed by the initials of the authors that first employed these two different envelopment surfaces (Banker, Charnes, & Cooper, 1984; Charnes et al., 1978).

[4] Flanders is the Dutch-speaking, northern part of the Belgian federal state.

[5] http://www.deafrontier.net/.

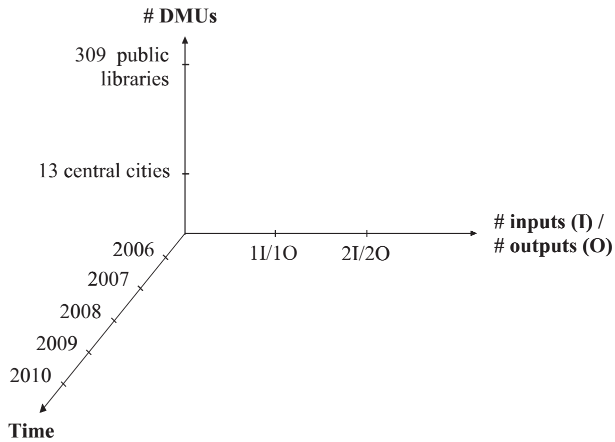


Fig. 3. Base scheme for efficiency benchmarking of Flemish public libraries.

Finally, the time aspect is reflected on the Z-axis. As previously mentioned, comparing an entity's own performance within time is also a type of benchmarking. For the case studies of public libraries in Flanders, there were data available in the Bios2 database from 2006 onwards.

The different expansion paths of efficiency benchmarking, applied to the public libraries in Flanders, will be presented. Starting with a single input, single output case of the Flemish central city libraries, the case studies examine how benchmarking can move along each of the three axes on the above-illustrated base scheme, as summarized in Table 2. Appendix A offers a representation of the four cases on the base figure.

5.2. Case 1: 13 Flemish central cities, year 2010, 1 input, 1 output

5.2.1. Procedures

The first case of benchmarking local public libraries in Flanders using the non-parametric frontier methods FDH and DEA involved a limited set of DMUs, a single efficiency measure, and a single year of analysis (2010; see Case 1 in Appendix A).

The limited set of comparable DMUs consisted of the 13 Flemish central city libraries. These 13 central cities are also considered as a set of comparable local authorities by the Agency of Internal Affairs of the Flemish Government and were already included as a cluster in the Flemish data portal, Local Statistics (Lokale Statistieken, 2011).

Concerning the selection of the single input and single output, the list of library efficiency studies in the literature review can be helpful. The overview in Table 1 reveals that the most relevant and often used inputs are staff or expenditure numbers, the main output candidate is circulation of materials. Based on these findings, and taking into account the available parameters in the Bios2 database, the following input and output for this first benchmarking case of local public libraries were selected:

- Input: expenditures (total expenditures during a calendar year);
- Output: circulation (total number of loans and renewals in a calendar year for the main library and all branch libraries).

Table 2
Overview of the four case studies.

Case	Entities (libraries in…)	Number of inputs – outputs	Time period
Case 1	13 central cities	1 input – 1 output	2010
Case 2	13 central cities	2 inputs – 2 outputs	2010
Case 3	79 residential LGs	1 input – 1 output	2010
Case 4	13 central cities	1 input – 1 output	2006–2010

Table 3
FDH efficiency scores, 1 input (expenditures) and 1 output (circulation).

Central cities	Input efficiency		Output efficiency		Dominating producer*
	Score	Rank	Score	Rank	
Aalst	0.778	11	0.826	10	Oostende/Oostende
Antwerpen	1.000	**1**	1.000	**1**	
Brugge	0.650	13	0.799	11	Leuven/Hasselt
Genk	0.829	10	0.673	13	Roeselare/Oostende
Gent	1.000	**1**	1.000	**1**	
Hasselt	1.000	**1**	1.000	**1**	
Kortrijk	0.921	8	0.934	8	Oostende/Oostende
Leuven	1.000	**1**	1.000	**1**	
Mechelen	1.000	**1**	1.000	**1**	
Oostende	1.000	**1**	1.000	**1**	
Roeselare	1.000	**1**	1.000	**1**	
Sint-Niklaas	0.669	12	0.692	12	Roeselare/Oostende
Turnhout	0.867	9	0.856	9	Mechelen/Mechelen
Average	*0.901*		*0.906*		

* In terms of input efficiency/output efficiency.

The selection of the input and output factor also covered the current budgetary situation by which local public libraries are pressured to do more with less.

5.2.2. Results

Concerning the efficiency of the public libraries in the 13 central cities, Table 3 presents the results of the FDH analysis using a single output (circulation) and a single input (expenditures) for the year 2010. The data for this analysis was extracted from the Bios2 database. In addition to the results in Table 3 below, Fig. 4 offers a graphic illustration of the FDH benchmarking.

From the FDH results it can be concluded that seven central cities were located on the efficiency frontier (i.e., a 100% score), meaning that the public libraries in these cities operated as fully efficient, relative to the other libraries in the set for the selected efficiency measure *circulation/expenditures*. Put differently, no other city libraries in the set achieved the same or more circulation with less or the same level of expenditures.

The cities with an efficiency score of less than 100% were found not to be fully efficient and therefore lie below the efficiency frontier (see Fig. 4). The further a city is located from the frontier, the larger its inefficiency (in relative terms). This inefficiency can be expressed in two ways. Take the example of Brugge:

a) Input inefficiency: Brugge has an FDH input efficiency score of 0.650 and is dominated by Leuven. Leuven needs less input to generate more output. The efficiency loss of Brugge is 0.350 of the total expenditures or, in other words, Brugge is only 65% input efficient.
b) Output efficiency: Brugge has an FDH output efficiency score of 0.799 and is dominated by Hasselt. Hasselt generates more output with less means. The efficiency loss of Brugge is 0.201 in terms of circulation, or in other words, Brugge is only 80% output efficient.

The same interpretation can also be made for the other FDH inefficient central city libraries. The scope for efficiency improvement was the largest for Sint-Niklaas and Brugge on the input side and for Sint-Niklaas on the output side (see Fig. 4).

The average FDH efficiency of the 13 central city public libraries was around 0.90, both for input and output efficiency. The input inefficiency implies that, on average, the DMUs in this set were able to achieve the same level of output (i.e., number of loans and renewals) with only 90% of their spending, meaning that they currently “waste” an average 10% of their financial resources. The output inefficiency implies that the cities generated only 90% of the technically possible output level with the current level of expenditures.

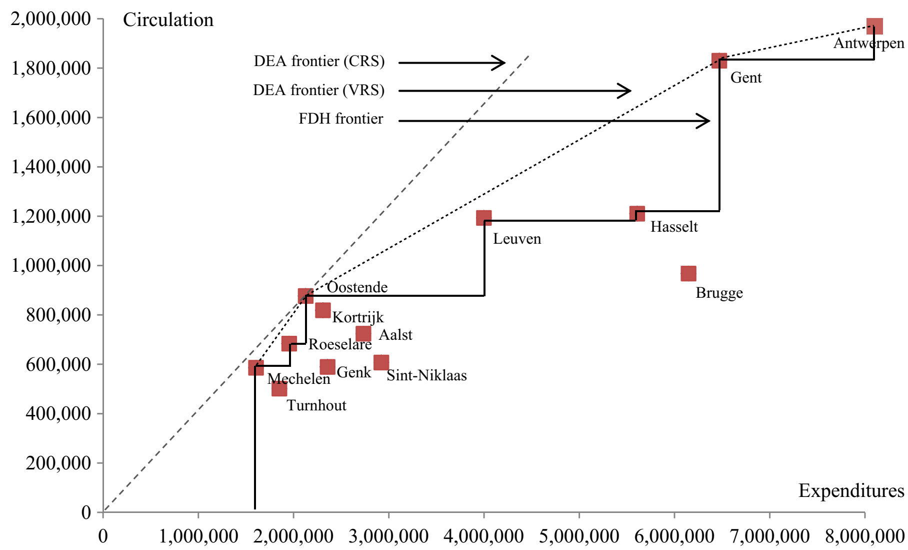


Fig. 4. FDH and DEA efficiency frontiers, 13 central cities, 2010, 1 input and 1 output.

In the FDH methodology and the accepted assumption therein, relative efficiency simply means the absence of observed better performance with the same or fewer resources, or observed allocation of fewer means with at least as large output. In a set with a relatively low number of observations (DMUs), one may lack the possibility of making many comparisons with similar ones. This induces a higher proportion of efficient DMUs. There is thus a *sparsity bias* in favour of those units that lie in a range where other observations are scarce. This bias leads to a significant overestimation of the number of efficient units and thus constitutes a serious shortcoming of the FDH approach (de Sousa & Schwengber, 2005).

An extreme form of the sparsity bias is the so-called *efficiency by default*. It refers to the following possible situations (Michailov, Tomova, & Nenkova, 2003):

- The library with the lowest level of spending and those with the highest value for at least one output by construction of the model will be declared 100% efficient. In other words, FDH will always indicate at least two fully efficient DMUs;
- Some observations, deemed efficient by not being dominated, do not dominate any other observation. Their efficiency is due to the absence of other observations with which the required comparison can be made.

In the FDH results in Table 3, the last column of dominating producers demonstrates the efficiency by default concept. Seven out of 13 central cities were efficient in the FDH model because no other central city library dominated them. Some of them dominated other inefficient DMUs; for instance, Oostende dominated Aalst and Kortrijk at the input side and Aalst, Genk, Kortrijk, and Sint-Niklaas at the output side. These cities are called *efficient by dominating*. A further investigation of the results reveals that two of the seven efficient cities (Antwerpen and Gent) did not dominate any other. Antwerpen and Gent were thus *efficient by default* [6] simply because they (1) were not dominated by any other DMU and (2) also did not dominate any other.

The DEA efficiency analysis scores are summarised in Table 4. As previously mentioned, DEA analysis differs from FDH by adding the assumption of frontier convexity. Convexity implies that if two observations are possible, then all their linear combinations are also possible. This leads to the fact that the DMUs in the sample (the 13 central city libraries) were not only compared relative to each other, but also with virtual DMUs (linear combinations of the existing fully efficient DMUs).

When the DEA benchmarking scores[7] were compared with those of the FDH analysis, only four central cities were found to be 100% efficient. Libraries found efficient under the DEA methodology were also efficient in the FDH analysis, but the inverse is not always true. Also, the average efficiency in the DEA analysis was lower (around 0.80) than in the FDH model. Contrary to the FDH benchmarking, Hasselt, Leuven, and Roeselare were no longer fully efficient in the DEA model, as can also be seen on Fig. 4. Hasselt and Leuven, for example, were 100% efficient with FDH because there was no other library that could achieve more or the same output with fewer means, but were inefficient with DEA in relation to Gent and Oostende because DEA does not only compare with Gent and Oostende, but also with all convex combinations of the input-output transformation of Gent and Oostende. In this way, the DEA model identifies for each DMU a (group of) possible benchmark(s).

It must be noted that such a graphic representation is only possible in a two-dimensional space, with one input and one output indicator.

5.3. Case 2: 13 Flemish central cities, year 2010, more inputs, more outputs

5.3.1. Procedures

In this second case, the efficiency benchmarking of public libraries was expanded by including more than one input and output in the analysis. Incorporating multiple inputs and outputs, either in *physical or financial* terms, is one of the major strengths of the non-parametric frontier methods. Inspiration for the selection of the multiple-input, multiple-output vector was again found in the literature review (see Table 1) and the available variables in the Bios2 database. The number of library staff (converted into FTEs) and the total collection expenditures (on all types of materials) were considered to be reliable input proxies. Both inputs were used to fulfil the main task of libraries (i.e., the loan of materials), taking the number of service hours of the

[6] Sometimes also expressed as ‘independent efficient’ (Gupta & Verhoeven, 2001).

[7] As a measure of comparison, the DEA scores under constant returns to scale (CSR) are additionally mentioned in Table 4; however only the results under variable returns to scale (VRS) are discussed, as it seems logical that the circulation (number of loans) will not increase proportionally with the resources spent but, rather, will increase in a diminishing proportion until a certain ceiling.

Table 4
DEA efficiency scores, 1 input (expenditures) and 1 output (circulation).

Central cities	Input oriented		Output oriented		Peers Input/Output	CRS
	VRS	Rank	VRS	Rank		
Aalst	0.678	10	0.717	9	Mechelen, Oostende/Gent, Oostende	0.643
Antwerpen	1.000	**1**	1.000	**1**	Antwerpen/Antwerpen	0.605
Brugge	0.414	13	0.550	13	Gent, Oostende/Gent, Oostende	0.382
Genk	0.684	9	0.636	11	Mechelen, Oostende/Gent, Oostende	0.607
Gent	1.000	**1**	1.000	**1**	Gent/Gent	0.687
Hasselt	0.650	11	0.738	8	Gent, Oostende/Gent, Oostende	0.524
Kortrijk	0.877	7	0.894	6	Mechelen, Oostende/Gent, Oostende	0.861
Leuven	0.893	6	0.927	5	Gent, Oostende/Gent, Oostende	0.724
Mechelen	1.000	**1**	1.000	**1**	Mechelen/Mechelen	0.886
Oostende	1.000	**1**	1.000	**1**	Oostende/Oostende	1.000
Roeselare	0.912	5	0.877	7	Mechelen, Oostende/Mechelen, Oostende	0.849
Sint-Niklaas	0.562	12	0.577	12	Mechelen, Oostende/Gent, Oostende	0.504
Turnhout	0.867	8	0.694	10	Mechelen/Mechelen, Oostende	0.658
Average	*0.811*		*0.816*			*0.687*

CRS – constant returns to scale.
VRS – variable returns to scale.

library into account. Only two inputs and two outputs were used in this case, due to the small number of DMUs (see Section 6). In summary:

- Input: (1) FTEs (personnel in FTEs); and (2) collection expenditures (for printed and audio-visual material, including serials);
- Output: (1) circulation (total number of loans and renewals in a calendar year for the main library and all branch libraries); and (2) opening hours (total weekly opening hours of the main library and all branches).

The set of DMUs and the year of analysis remained the same as in the first case, meaning that the public libraries in the 13 Flemish central cities were benchmarked for the year 2010. Compared with the first, this case moves along the X-axis on the base scheme in Fig. 3, which is also illustrated in Appendix A. Again, the data for this analysis was extracted from the Bios2 database.

5.3.2. Results

Table 5 presents the FDH efficiency benchmarking for two inputs and two outputs, where it can be seen that only one central city library was scored relatively inefficient. In this way, the FDH model with these features can be considered as irrelevant. Nevertheless, the results are included to illustrate the efficiency by default phenomenon and to stress the necessary requirements when applying the FDH technique.

Table 5
FDH, 2 inputs (FTEs, collection expenditures) and 2 outputs (opening hours, circulation).

Central city	Input efficiency		Output efficiency		Dominating producer*
	Score	Rank	Score	Rank	
Aalst	1.000	**1**	1.000	**1**	
Antwerpen	1.000	**1**	1.000	**1**	
Brugge	1.000	**1**	1.000	**1**	
Genk	1.000	**1**	1.000	**1**	
Gent	1.000	**1**	1.000	**1**	
Hasselt	1.000	**1**	1.000	**1**	
Kortrijk	1.000	**1**	1.000	**1**	
Leuven	1.000	**1**	1.000	**1**	
Mechelen	1.000	**1**	1.000	**1**	
Oostende	1.000	**1**	1.000	**1**	
Roeselare	1.000	**1**	1.000	**1**	
Sint-Niklaas	1.000	**1**	1.000	**1**	
Turnhout	0.737	13	0.856	13	Mechelen/ Mechelen
Average	*0.980*		*0.989*		

* In terms of input efficiency/output efficiency.

The fact that nearly all DMUs were scored 100% efficient in this multiple-input, multiple-output FDH analysis, can be explained by the occurrence of efficiency by default when a small set of DMUs is taken into account (see above). Of all 12 efficient observations, only Mechelen was efficient by dominating (namely dominating the inefficient DMU Turnhout). All other 11 efficient libraries were efficient by default. Because the FDH technique only performs comparisons with other existing DMUs in the set (not with linear combinations as in DEA), the small set of observations leads to a shortage of mutual comparisons that can be made. In other words, with this limited number of DMUs, the FDH model with two inputs and two outputs is not able to classify units as efficient or inefficient. To avoid this overestimation of efficient units, the number of observations must increase exponentially as the number of inputs and outputs increase (Simar & Wilson, 2000).

When the relative efficiency of the 13 public libraries was analysed using DEA (see Table 6), it became apparent that there were significantly fewer fully efficient DMUs, because DEA also made domination comparisons with virtual DMUs. Five central city libraries were not dominated by a peer or linear combinations of peers, and therefore scored 100% efficient in their transformation of the two inputs into the two outputs. In respect to the previous single-input, single-output case, Leuven and Roeselare became efficient, contrary to Gent, which scored as inefficient. Antwerpen, Mechelen, and Oostende remained efficient in comparison to Case 1.

The other eight central cities were not on the DEA efficiency frontier, and thus had a score of less than 1. Compared to the DEA analysis in case 1, Hasselt and Turnhout were the main laggards and Brugge and Sint-Niklaas performed slightly better. Hasselt, for instance, was dominated by (linear combinations of) Antwerpen, Leuven, and Mechelen in terms of input. Leuven was able to generate nearly the same output (circulation and opening hours) as Hasselt with less than half staff numbers (FTEs) and collection expenses. This means that Hasselt wasted half its resources (i.e., nearly 50% input efficient). Turnhout, as another example, put in around the same amount of means (FTEs and collection expenditures) as Oostende, but achieved significantly lower output in comparison with Oostende. Turnhout was thus found considerably (36%) output inefficient (i.e., output efficiency score of 64%).

Finally, this second case of multiple inputs and outputs makes clear that more observations must be included to make an FDH efficiency analysis relevant. The DEA analysis does not differ significantly with the single-input, single-output case, and forms one step towards identifying sustainably efficient DMUs (i.e., DMUs that are scored efficient no matter what model is used).

Table 6
DEA, 2 inputs (FTEs, collection expenditures) and 2 outputs (opening hours, circulation).

Central city	Input oriented		Output oriented		Peers Input/Output	CRS
	VRS	Rank	VRS	Rank		
Aalst	0.743	10	0.712	11	Leuven, Mechelen/ Antwerpen, Leuven, Mechelen	0.708
Antwerpen	1.000	**1**	1.000	**1**	Antwerpen/Antwerpen	1.000
Brugge	0.750	9	0.772	10	Antwerpen, Mechelen, Roeselare/Antwerpen, Mechelen, Roeselare	0.743
Genk	0.730	12	0.803	9	Mechelen, Roeselare/ Antwerpen, Mechelen, Roeselare	0.722
Gent	0.902	6	0.926	6	Antwerpen, Leuven/Antwerpen, Leuven	0.754
Hasselt	0.532	13	0.657	12	Antwerpen, Leuven, Mechelen/Antwerpen, Leuven	0.499
Kortrijk	0.877	8	0.908	7	Antwerpen, Leuven, Mechelen/Antwerpen, Leuven, Mechelen	0.838
Leuven	1.000	**1**	1.000	**1**	Leuven/Leuven	1.000
Mechelen	1.000	**1**	1.000	**1**	Mechelen/Mechelen	1.000
Oostende	1.000	**1**	1.000	**1**	Oostende/Oostende	1.000
Roeselare	1.000	**1**	1.000	**1**	Roeselare/Roeselare	0.924
Sint-Niklaas	0.878	7	0.878	8	Antwerpen, Mechelen/ Antwerpen, Mechelen	0.877
Turnhout	0.737	11	0.639	13	Mechelen/Leuven, Mechelen, Oostende	0.630
Average	*0.858*		*0.869*			*0.823*

5.4. Case 3: 79 residential municipalities, year 2010, 1 input, 1 output

5.4.1. Procedures

The third case returns to a single input, single output efficiency analysis for the year 2010, but the number of libraries in the benchmarking model was enlarged. Instead of working with a limited set of DMUs (13 central cities), the analysis was repeated for a broad set of units. Compared to Case 1, the benchmarking in this third case extends along the Y-axis on the base scheme in Fig. 3, as also illustrated in Appendix A.

Again, attention must be paid to the comparability of the entities included in the analysis. For benchmarking to be effective, it is crucial that the compared organizations are as similar as possible in relevant external, environmental variables (Berg, 2010). Therefore, because of the large variance in types of local governments (e.g., in terms of scale or demographic profile), it was not appropriate to integrate all 309 Flemish public libraries in one benchmarking study.

For selecting an appropriate, but larger, set of comparable public libraries in this case, the classification of Flemish local governments developed by Belfius was used. This classification aims to categorize local governments with comparable socio-economic conditions into homogeneous clusters (Belfius, 2007). In this third case, the public libraries in the 79 Flemish *residential municipalities* were chosen as a set of DMUs.[8]

5.4.2. Results

Table 7 presents the results of both the FDH and DEA efficiency benchmarking of the public libraries in the 79 residential local governments in Flanders for 2010, using yearly expenditures as input- and circulation as output-indicators. The data for this analysis was extracted from the Bios2 database. As this case, again, covered a two dimensional space with one input and one output, a graphic representation of the FDH and DEA efficiency frontiers is possible and given in Fig. 5.

The results of the non-parametric efficiency analysis show that only a few residential local governments (LGs) were rated fully efficient. The FDH model indicated nine municipalities whose public library worked 100% efficiently: Boechout, Dilbeek, Gavere, Grimbergen, Herent, Huldenberg, Merelbeke, Wommelgem, and Zingem. Hence, these LGs construct the FDH efficiency frontier (see Fig. 5). A subgroup of three (Dilbeek, Huldenberg, and Wommelgem) remained fully efficient when applying the DEA model (under the variable returns to scale assumption).

Municipalities with an efficiency score below 1 faced a certain level of inefficiency, compared with the 100% efficient residential LG libraries (FDH), or additionally compared with linear combinations of these efficient LGs (DEA); meaning that there are other DMUs (even virtual ones in the DEA analysis) that are able to reach better performance (higher circulation) with a lower level of resources (less expenditures). The input inefficiency (i.e., 1 minus the efficiency score) indicates the level of "waste" in spending, whereas an output efficiency score below 1 means that other DMUs achieved higher circulation with the same or less input (see Case 1 for a more detailed description of the interpretation of the efficiency scores).

It is remarkable that there were proportionally fewer fully efficient DMUs in this exercise with 79 residential LG libraries (11.4% with FDH, 3.8% with DEA) than in the analysis of 13 central city libraries (53.8% with FDH, 30.8% with DEA). Also, the average efficiency of all DMUs was significantly lower than in the previous cases. Indeed, the more observations included in the sample, the better the approximation of the true frontier. This is due mainly to the larger number of observations, with the consequence that fewer DMUs will be found efficient by default. In this case, only Huldenberg was efficient by default.

As stated earlier, the relevance of benchmarking with non-parametric frontier methods lies in the fact that the benchmarked entities can judge their performance relative to other comparable units. In this case, a public library such as DMU *A* in Fig. 5 is so deviated from the constructed efficiency frontier that it will be useful for such entities to investigate further what factors may cause their relative inefficiency (i.e., benchlearning). DMU *A* can identify Dilbeek as a *best practice* (same expenditure level, but significantly more output) from which to learn. On the other hand, entities such as library *B* in Fig. 5 will be less alarmed by this benchmarking exercise.

5.5. Case 4: 13 Flemish central cities, years 2006–2010, 1 input, 1 output

5.5.1. Procedures

As previously mentioned, benchmarking involves not only comparing organizational performance to peers in the area, but also assessing how the relative performance of organizations changes over a certain time period. This fourth and last case returns to the public libraries in the 13 central cities, but benchmarks their relative efficiency in time, based on one input (expenditures) and one output (circulation) for a

[8] The Belfius classification defines the residential municipalities as a main socio-economic category, consisting of clusters v1, v2, v10 and, v11. The complete classification can be found at Belfius (2007).

Table 7
FDH and DEA efficiency scores, 79 residential LG, 2010, 1 input (expenditures) and 1 output (circulation).

Residential LG	Input efficiency		Output efficiency	
	FDH	DEA-VRS	FDH	DEA-VRS
Aartselaar	0.235	0.232	0.648	0.550
Affligem	0.900	0.687	0.209	0.209
Asse	0.629	0.421	0.806	0.662
Beersel	0.180	0.172	0.442	0.442
Begijnendijk	0.500	0.398	0.284	0.239
Bertem	0.764	0.604	0.463	0.432
Bierbeek	0.687	0.627	0.775	0.701
Boechout	1.000	0.897	1.000	0.952
Bonheiden	0.365	0.329	0.609	0.528
Borsbeek	0.542	0.435	0.493	0.412
Boutersem	0.691	0.536	0.424	0.384
Brasschaat	0.823	0.802	0.984	0.984
Buggenhout	0.436	0.427	0.776	0.727
De Pinte	0.839	0.666	0.468	0.448
Destelbergen	0.466	0.399	0.630	0.498
Dilbeek	1.000	1.000	1.000	1.000
Edegem	0.264	0.256	0.665	0.557
Erpe-Mere	0.385	0.346	0.605	0.537
Gavere	1.000	0.765	1.000	0.890
Grimbergen	1.000	0.734	1.000	0.851
Grobbendonk	0.547	0.428	0.443	0.371
Haacht-Boortmeerbeek	0.446	0.243	0.763	0.544
Herent	1.000	0.863	1.000	0.945
Hoeilaart	0.908	0.751	0.553	0.541
Holsbeek	0.714	0.709	0.983	0.899
Hove	0.582	0.453	0.430	0.369
Huldenberg	1.000	1.000	1.000	1.000
Jabbeke	0.488	0.438	0.738	0.593
Kalmthout	0.683	0.552	0.862	0.777
Kampenhout	0.680	0.527	0.420	0.378
Kapellen	0.843	0.574	0.953	0.743
Kapelle-op-den-Bos	0.932	0.713	0.394	0.388
Keerbergen	0.430	0.366	0.506	0.471
Kontich	0.677	0.347	0.750	0.673
Kortenberg	0.649	0.292	0.727	0.639
Lennik	0.961	0.732	0.525	0.239
Lint	0.752	0.603	0.301	0.287
Londerzeel	0.347	0.309	0.594	0.503
Lovendegem	0.772	0.610	0.461	0.432
Lubbeek	0.409	0.371	0.628	0.572
Machelen	0.493	0.395	0.490	0.396
Meise	0.349	0.279	0.398	0.338
Melle	0.537	0.490	0.774	0.645
Merchtem	0.311	0.272	0.489	0.445
Merelbeke	1.000	0.704	1.000	0.829
Mortsel	0.651	0.452	0.819	0.720
Nazareth	0.998	0.568	0.874	0.825
Nevele	0.587	0.494	0.593	0.509
Nijlen	0.399	0.367	0.652	0.588
Oosterzele	0.340	0.331	0.678	0.644
Opwijk	0.725	0.594	0.340	0.321
Oud-Heverlee	0.402	0.338	0.488	0.441
Oud-Turnhout	0.615	0.503	0.529	0.462
Overijse	0.414	0.319	0.340	0.311
Putte	0.464	0.365	0.454	0.358
Ranst	0.529	0.511	0.914	0.757
Roosdaal	0.971	0.766	0.456	0.453
Rotselaar	0.513	0.445	0.660	0.541
Schilde	0.350	0.342	0.770	0.655
Schoten	0.849	0.608	0.824	0.824
Sint-Katelijne-Waver	0.793	0.424	0.859	0.735
Sint-Lievens-Houtem	0.658	0.526	0.486	0.434
Sint-Martens-Latem	0.697	0.537	0.409	0.371
Sint-Pieters-Leeuw	0.646	0.511	0.856	0.750
Steenokkerzeel	0.616	0.490	0.476	0.416
Ternat	0.681	0.551	0.507	0.458
Tervuren	0.307	0.283	0.579	0.524
Tremelo	0.549	0.470	0.631	0.529
Vosselaar	0.697	0.546	0.444	0.403
Waasmunster	0.695	0.582	0.383	0.357
Wemmel	0.577	0.464	0.306	0.270
Wijnegem	0.872	0.707	0.318	0.315
Wommelgem	1.000	1.000	1.000	1.000

(continued on next page)

Table 7 *(continued)*

Residential LG	Input efficiency		Output efficiency	
	FDH	DEA-VRS	FDH	DEA-VRS
Zandhoven	0.407	0.325	0.400	0.364
Zaventem	0.265	0.257	0.662	0.556
Zemst	0.378	0.365	0.753	0.663
Zingem	1.000	0.840	1.000	0.510
Zoersel	0.489	0.225	0.666	0.550
Average	*0.640*	*0.518*	*0.642*	*0.564*

time range 2006–2010. Compared to Case 1, this exercise moves thus along the Z-axis on the base scheme in Fig. 3, as illustrated in Appendix A.

5.5.2. Results

First, the 13 central city libraries were benchmarked in the same way as in Case 1 and repeated for the other years 2006–2009. Again, the data for this analysis were extracted from the Bios2 database. Table 8 presents both the FDH and DEA efficiency benchmarking scores. Of note, Antwerpen, Gent, and Mechelen stayed fully efficient during the whole time range. For Antwerpen, this was due mainly to the fact that it always had the largest output and thus could be dominated by others in the set (efficiency by default). Similarly, Turnhout was 100% efficient in 2006–2009 because it had the lowest level of expenses. For the other DMUs, the efficiency scores varied from year to year. In this manner, the time path concerns a measurement of the relative (in)efficiency of the central city libraries over time, and can force cities to analyse further the differences with their counterparts in a certain year or to explain fluctuations of scores over time.

In a second way, for more in-depth insight in the relative efficiency, the evolution of input (expenditures) and output (circulation) can be illustrated graphically for each of the 13 central city libraries, as done in Fig. 6. A good example in this case is the time path for the library in Roeselare, where a substantial rise in expenditures towards 2008 was evident, but thereafter fell back to its initial position of 2006. As there were demonstrable reasons for the great increase in expenses in 2008 (e.g., a specific investment) the large inefficiency in 2008 could probably be sufficiently explained. The same reasoning holds for Genk in the year 2007 and 2008, Hasselt in 2008, and Brugge in 2010.

The illustrated time path can also be useful for other cities to analyse their relative efficiency results more in detail. Mechelen, for example, was faced with declining output while holding the input level constant. This means that Mechelen's own efficiency was decreasing over time, although it was found to be fully efficient relative to the other central cities in all the years. The same conclusion could be made for Turnhout and Sint-Niklaas was also confronted with decreasing output and increasing expenses. The latter case was an indication of loss in efficiency over time, although its relative efficiency (compared to the other city libraries) remained fairly constant. A reverse tendency existed for Aalst in 2009: although the expenses hardly increased, this library was faced with a strong growth in circulation; its efficiency thus rose but, compared to the other central cities, there remained a large level of inefficiency.

6. Discussion

The results from this empirical work in benchmarking the efficiency of public libraries allow several conclusions. Concerning the integration of benchmarking in time and in space, both ways are proven to have merits. By comparing efficiency over time, organizations can reveal success factors in times of increasing efficiency or detect reasons for declining performance. On the other hand, benchmarking against similar entities is also important, as increasing performance over time does not necessarily or technically mean that the organization performs efficiently, relative to other organisations in the field.

Concerning the use of the two non-parametric frontier methods, three conclusions can be drawn. First, the results using DEA are stricter in respect to results using FDH, meaning that an entity that is scored as efficient under DEA is also efficient under FDH, although the reverse is not true. Second, a distinguishing feature of the FDH/DEA methodology is its capability to handle multiple inputs and outputs at once; but when working in such multi-vector space, sufficient observations must to be included in the analysis to avoid overestimation due to the efficiency by default phenomenon. Third, the presented cases confirm that FDH

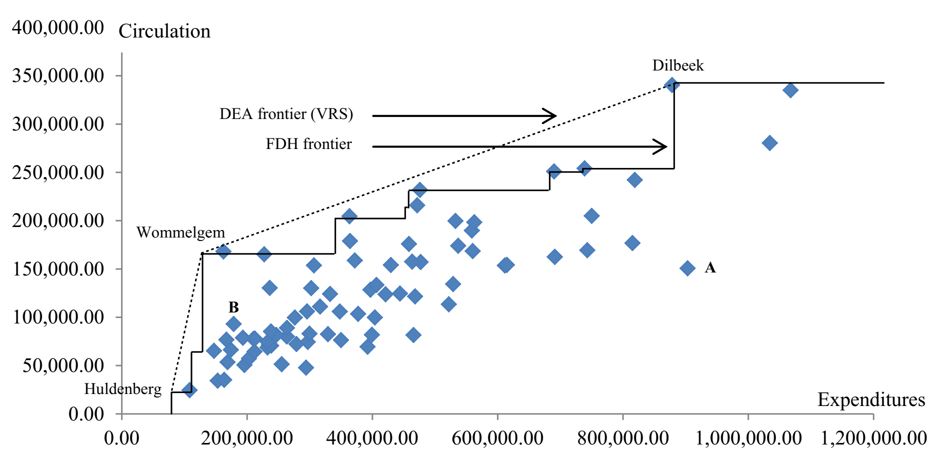


Fig. 5. Efficiency frontier FDH and DEA, 79 rural LGs, 2010, 1 input and 1 output.

Table 8
FDH and DEA efficiency scores, 1 input (expenditures) and 1 output (circulation), 2006–2010.

Central city_score in year x	Input efficiency		Output efficiency	
	FDH	DEA-VRS	FDH	DEA-VRS
Aalst_2006	0.642	0.642	0.764	0.608
Aalst_2007	0.666	0.666	0.697	0.632
Aalst_2008	0.579	0.579	0.588	0.499
Aalst_2009	0.766	0.663	0.810	0.684
Aalst_2010	0.778	0.678	0.826	0.717
Antwerpen_2006	1.000	1.000	1.000	1.000
Antwerpen_2007	1.000	1.000	1.000	1.000
Antwerpen_2008	1.000	1.000	1.000	1.000
Antwerpen_2009	1.000	1.000	1.000	1.000
Antwerpen_2010	1.000	1.000	1.000	1.000
Brugge_2006	1.000	0.685	1.000	0.731
Brugge_2007	0.818	0.705	0.903	0.745
Brugge_2008	0.692	0.594	0.910	0.691
Brugge_2009	0.768	0.638	0.892	0.719
Brugge_2010	0.650	0.414	0.799	0.550
Genk_2006	0.775	0.775	0.739	0.653
Genk_2007	0.477	0.477	0.568	0.424
Genk_2008	0.424	0.424	0.394	0.385
Genk_2009	0.719	0.696	0.682	0.633
Genk_2010	0.829	0.684	0.673	0.636
Gent_2006	1.000	1.000	1.000	1.000
Gent_2007	1.000	1.000	1.000	1.000
Gent_2008	1.000	1.000	1.000	1.000
Gent_2009	1.000	1.000	1.000	1.000
Gent_2010	1.000	1.000	1.000	1.000
Hasselt_2006	0.952	0.646	0.993	0.696
Hasselt_2007	1.000	0.893	1.000	0.910
Hasselt_2008	0.652	0.591	0.943	0.683
Hasselt_2009	1.000	0.678	1.000	0.743
Hasselt_2010	1.000	0.650	1.000	0.738
Kortrijk_2006	1.000	1.000	1.000	1.000
Kortrijk_2007	1.000	0.890	1.000	0.917
Kortrijk_2008	0.838	0.699	0.812	0.733
Kortrijk_2009	0.840	0.756	0.859	0.770
Kortrijk_2010	0.921	0.877	0.934	0.894
Leuven_2006	0.406	0.374	0.741	0.442
Leuven_2007	1.000	0.813	1.000	0.839
Leuven_2008	1.000	0.988	1.000	0.992
Leuven_2009	1.000	0.977	1.000	0.983
Leuven_2010	1.000	0.893	1.000	0.927
Mechelen_2006	1.000	1.000	1.000	1.000
Mechelen_2007	1.000	1.000	1.000	1.000
Mechelen_2008	1.000	1.000	1.000	1.000
Mechelen_2009	1.000	1.000	1.000	1.000
Mechelen_2010	1.000	1.000	1.000	1.000
Oostende_2006	0.901	0.764	0.858	0.792
Oostende_2007	1.000	0.929	1.000	0.946
Oostende_2008	1.000	1.000	1.000	1.000
Oostende_2009	1.000	1.000	1.000	1.000
Oostende_2010	1.000	1.000	1.000	1.000
Roeselare_2006	1.000	0.852	1.000	0.862
Roeselare_2007	0.562	0.435	0.604	0.522
Roeselare_2008	0.218	0.216	0.336	0.321
Roeselare_2009	0.764	0.645	0.779	0.657
Roeselare_2010	1.000	0.912	1.000	0.877
Sint-Niklaas_2006	0.770	0.653	0.780	0.690
Sint-Niklaas_2007	0.597	0.596	0.806	0.680
Sint-Niklaas_2008	0.577	0.571	0.745	0.611
Sint-Niklaas_2009	0.750	0.609	0.734	0.611
Sint-Niklaas_2010	0.669	0.562	0.692	0.577
Turnhout_2006	1.000	1.000	1.000	1.000
Turnhout_2007	1.000	1.000	1.000	1.000
Turnhout_2008	1.000	1.000	1.000	1.000
Turnhout_2009	1.000	1.000	1.000	1.000
Turnhout_2010	0.867	0.867	0.856	0.694

and DEA methodologies tend to give significantly different results. A combination of both techniques could therefore be of great value to both managerial and strategic decision-makers. On the one hand, FDH analysis is convincing as it identifies the most obvious and real counterparts for the inefficient DMUs to learn from in terms of realistic comparable best practices. On the other hand, DEA has the advantage of providing efficient goals for DMUs to work towards: the basis for benchlearning; although these goals should be subject to further study in terms of their feasibility in practice.

In applying these FDH and DEA techniques, some considerations must be addressed. First, it is clear that frontier methods form a data-driven approach of performance measurement, meaning that the necessary data quality requirements have to be met. Non-parametric frontier methods compare entities (DMUs) relatively, against each other, whereby the best performing organizations (i.e., those with the highest output-input mix) construct a frontier which envelops the set of DMUs. The inefficiency of entities can then be measured as the radial distance of a DMU to the frontier (the further from the efficiency frontier, the larger the inefficiency) (Tulkens, 1993). Herein lies one of the limitations of these deterministic frontier techniques, as each deviation from the frontier – even due to random noise or measurement error – is considered as inefficiency. This, of course, means that the results of such benchmarking are highly sensitive for extreme observations in the set (*outliers*) and for model specifications in case of a small sample (Pedraja-Chaparro, Salinas-Jiménez, & Smith, 2005). Nevertheless, more recently both highly complex statistical tests and bootstrapping methods have been developed for dealing with this issue (see e.g., Banker & Natarajan, 2004; Simar & Wilson, 2008).

Moreover, the inclusion or exclusion of variables can influence the efficiency estimates. The exclusion of an important input or output that is not correlated with variables included in the analysis can bias the results (i.e., underestimate the efficiency), just as the inclusion of an irrelevant input or output can lead to an efficiency overestimation (Smith, 1997; Thanassoulis, 2001). The selection of inputs and outputs – a choice made by the analyst – therefore must be well considered. Also, the number of variables included in the model is important, especially when a limited set of DMUs is taken into account. A general rule of thumb is that there should be at least three times as many DMUs as there are inputs and outputs combined (Jacobs, Smith, & Street, 2006). Besides that, the more variables included in the analysis, the less discriminating the model tends to be. Therefore, the second case with multiple inputs and outputs counted more fully efficient DMUs compared to the single input-output model.

Two final considerations that must be made when using non-parametric frontier methods include choosing between input or output orientation and (only with DEA) assuming constant or variable returns of scale. Again, these trade-offs need to be considered by the analyst. Input and output oriented models will determine the same set of fully efficient DMUs (the same frontier), but the efficiency scores of inefficient entities will differ (Coelli, Rao, & Battese, 1998). Obviously, both types of scores can be reported, as has been done in this study. But it is possible that decision-makers or managers can only influence input parameters (e.g., personnel or expenditures) and not the output produced with these resources. In that case, input orientation seems most appropriate to analyse which proportional reduction in resources is achievable given a certain level of output. Contrarily, when the analyst is interested in maximizing output given the level of staffing, an output-oriented model will have to be selected. Also, the choice of constant (CRS) or variable (VRS) returns to scale, in the case of applying DEA, will usually depend on the context and purpose of the analysis. From a social perspective, interest may be in productivity regardless of the scale of operations, so CRS may be more appropriate. From a managerial perspective, the focus will be on the extent to which the scale of operations affects productivity, so VRS may be preferred (Jacobs et al., 2006).

The exploratory nature of this study must be stressed. Based on the prevailing findings, there are several directions for further research. First of all, more in-depth analysis is required to determine if the measured scores reflect genuine inefficiencies or if they are explained by the action of others factors. For instance, in some

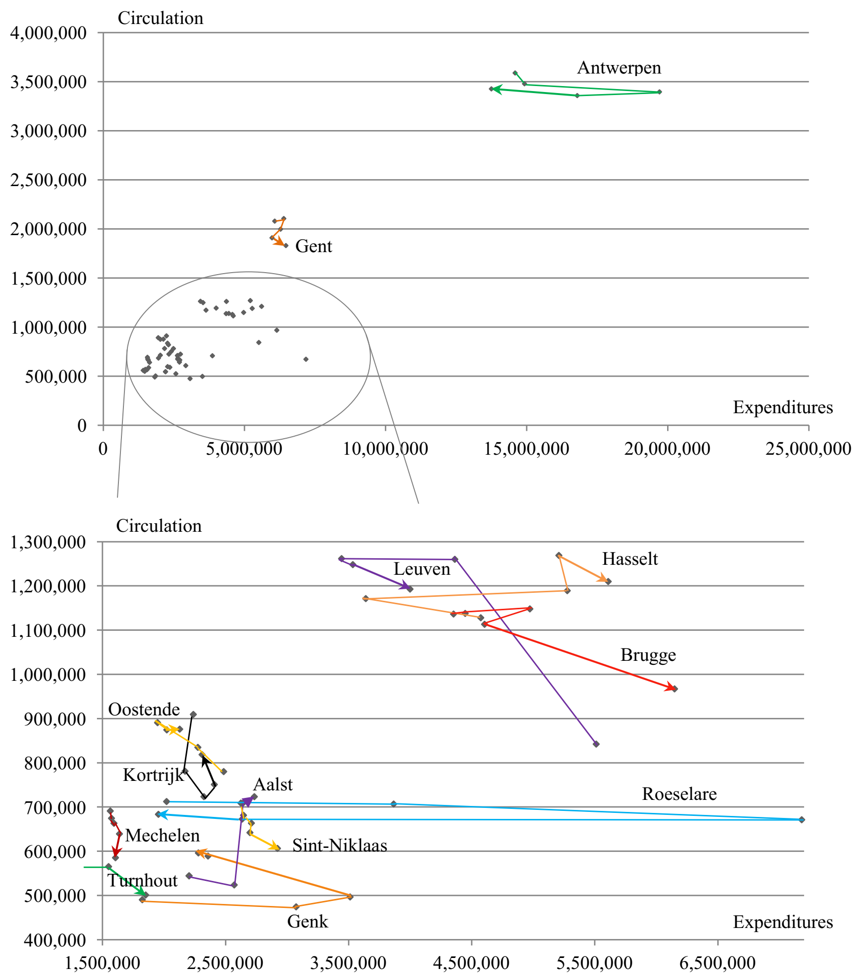


Fig. 6. 13 central cities, time path for 1 input and 1 output (2006–2010).

cases, inefficient libraries may be intrinsically different from the dominating ones and what is regarded as inefficiency could correspond simply to the effects of such library-specific characteristics. In particular, no attempt was made to include variables reflecting the quality of the services provided by libraries. Second, as efficiency evaluation is highly dependent on input and output selection, more detailed analysis of potential input and output candidates can help increase both the validity and reliability of the benchmarking process. Third, future work can assess the relation between the benchmarking results and organisational learning. This relationship is often referred to as benchlearning (Karlöf & Östblom, 1995). Having conducted benchmarking, the superior performers (i.e., the 100% efficient libraries) are identified and additional comparison of how processes are designed and implemented by those best performing entities will shape the organisational learning potential for the purpose of performance improvement.

7. Conclusion

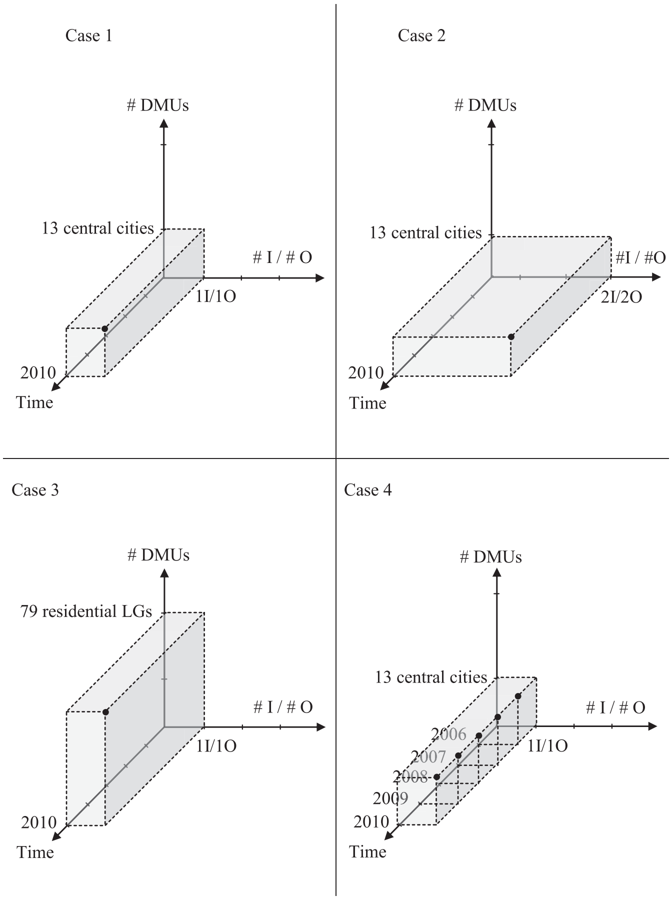
Benchmarking is widely acknowledged to be a useful instrument in assessing and improving organizational performance. More recently, the use of non-parametric frontier methods is introduced as an alternative benchmarking approach instead of comparing a set of indicators, which allow only a partial evaluation of performance. The present study demonstrates that combining the two non-parametric frontier techniques, FDH and DEA, gives supplementary results and is therefore of great value for organizations in identifying relevant benchmarks and providing performance objectives.

Moreover, an innovative approach to non-parametric efficiency benchmarking is presented in this study by incorporating all possible paths of expansion into a coherent whole. In this way, an integrated efficiency analysis – a combination of benchmarking over time and against peers – can perfectly build a bridge towards benchlearning and efficiency improvements trajectories. Entities that are found to be fully efficient can evaluate whether their efficiency is sustainable over longer time periods and those found to be inefficient can identify counterparts (best practices in the field) from which they can learn. The holistic approach used in the present study can be applied to any other group of public sector institutions where questions about efficiency and performance improvement have arisen.

Acknowledgments

We thank the participants of Workshop 3 at the 9th Transatlantic Dialogue (Balitmore, MD, USA), in particular Fran Berry from Florida State University, for their helpful comments and suggestions. We also thank the Policy Research Centre on Governmental Organization – Decisive Government (SBOV III), for funding this research.

Appendix A. Illustration of the four case studies.



References

Afonso, A., & St. Aubyn, M. (2005). Non-parametric approaches to education and health efficiency in OECD countries. *Journal of Applied Economics*, *8*(2), 227–246.

Agasisti, T., & Johnes, G. (2009). Beyond frontiers: Comparing the efficiency of higher education decision-making units across more than one country. *Education Economics*, *17*(1), 59–79.

Ammons, D. (2012). *Municipal benchmarks: Assessing local performance and establishing community standards.* New York, NY: ME Sharpe.

Askim, J., Johnsen, Å., & Christophersen, K. -A. (2008). Factors behind organizational learning from benchmarking: Experiences from Norwegian municipal benchmarking networks. *Journal of Public Administration Research and Theory*, *18*(2), 297–320.

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, *30*(9), 1078–1092.

Banker, R. D., & Natarajan, R. (2004). Statistical tests based on DEA efficiency scores. In W. W. Cooper, L. Seiford, & J. Zhu (Eds.), *Handbook on data envelopment analysis* (pp. 299–321). Norwell, MA: Kluwer Academic.

Belfius (2007). *Problematiek van de classificatie van de gemeenten.* Retrieved from https://www.belfius.be/publicsocial/nl/media/typologie_new_nl_tcm_31-36262.pdf

Belfius (2013). *Fiche 1: De macro-economische context. Fiches over de financiële uitdagingen voor de gemeenten.* Retrieved from https://www.belfius.be/publicsocial/nl/expertise/studies/lokalefinancien/financieleuitdagingen/index.aspx

Berg, S. V. (2010). *Water utility benchmarking: Measurement, methodologies, and performance incentives.* London, UK: IWA Publishing.

Bios2 (2013). *Bibliotheek informatie-en opvolgingssyteem.* Retrieved from http://www.bibliotheekstatistieken.be/

Bourgeois, G. (2009). *Beleidsnota binnenlands bestuur 2009–2014.* Brussels, Belgium: Vlaamse overheid.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, *2*(6), 429–444.

Chen, T. (1997). An evaluation of the relative performance of university libraries in Taipei. *Library Review*, *46*(3), 190–200.

Coelli, T., Rao, D., & Battese, G. (1998). *An introduction to efficiency and productivity analysis.* Norwell, MA: Kluwer Academic Publishers.

Cummins, J. D., & Weiss, M. A. (2013). Analyzing firm performance in the insurance industry using frontier efficiency and productivity models. In G. Dionne (Ed.), *Handbook of insurance* (pp. 795–861). New York, NY: Springer.

De Borger, B., & Kerstens, K. (1996). Cost efficiency of Belgian local governments: A comparative analysis of FDH, DEA, and econometric approaches. *Regional Science and Urban Economics*, *26*(2), 145–170.

de Carvalho, F. A., Jorge, M. J., Jorge, M. F., Russo, M., & de Sá, N. O. (2012). Library performance management in Rio de Janeiro, Brazil: Applying DEA to a sample of university libraries in 2006–2007. *Library Management*, *33*(4/5), 297–306.

de Sousa, M., & Schwengber, S. (2005, December). Efficiency estimates for judicial services in Brazil: Nonparametric FDH (free disposal hull) and the expected order-M efficiency scores for Rio Grande do Sul courts. *Paper presented at ANPEC 2005: 33th Encontro Nacional de Economia, 6–9 December, Natal, Brazil.*

De Witte, K., & Geys, B. (2011). Evaluating efficient public good provision: Theory and evidence from a generalised conditional efficiency model for public libraries. *Journal of Urban Economics*, *69*(3), 319–327.

De Witte, K., & Geys, B. (2012). *Citizen coproduction and efficient public good provision: Theory and evidence from local public libraries* (Discussion paper). Berlin, Germany: Social Science Research Center Berlin (WZB).

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of Royal Statistical Society: Series A*, *120*(3), 253–290.

Fenna, A. (2012). Benchmarking in federal systems. In A. Fenna, & F. Knüpling (Eds.), *Benchmarking in federal systems: Roundtable proceedings* (pp. 11–38). Canberra, Australia: Productivity Commission.

Gupta, S., & Verhoeven, M. (2001). The efficiency of government expenditure: Experiences from Africa. *Journal of Policy Modeling*, *23*(4), 433–467.

Hammond, C. (2002). Efficiency in the provision of public services: A data envelopment analysis of UK public library systems. *Applied Economics*, *34*(5), 649–657.

Herrera, S., & Pang, G. (2005). *Efficiency of public spending in developing countries: An efficiency frontier approach* (Policy research working paper no. 3645). Washington, DC: World Bank.

Jacobs, R., Smith, P. C., & Street, A. (2006). *Measuring efficiency in health care: Analytic techniques and health policy.* Cambridge, UK: Cambridge University Press.

Jo, S., Park, T., Lee, G., & Yoon, D. (2009). An analysis of the relative efficiency for the national university libraries using DEA model. *The Journal of Korea Library Information Science Society*, *40*(1), 253–274.

Kalb, A., Geys, B., & Heinemann, F. (2012). Value for money? German local government efficiency in a comparative perspective. *Applied Economics*, *44*(2), 201–218.

Karlöf, B., & Östblom, S. (1995). *Benchmarking: Tuottavuudella ja laadulla mestariksi.* Helsinki, Finland: Gummerus.

Kwack, Y. J. (1993). Research note: Analysis of efficiency trend in university's libraries: A data envelopment analysis approach. *Journal of Management*, *9*(1), 183–206.

Lokale Statistieken (2011). *13 centrumsteden.* Retrieved from http://aps.vlaanderen.be/lokaal/beleidsplannen/13steden.html

Mann, G. M. (1997). *Efficiency evaluations of North American university libraries: Data envelopment analysis.* Montreal, Canada: McGill University.

Michailov, A., Tomova, M., & Nenkova, P. (2003). Cost efficiency in Bulgarian municipalities. *Working paper s* (pp. 39–45). Sofia, Bulgaria: University for the National and World Economy.

Miidla, P., & Kikas, K. (2009). The efficiency of Estonian central public libraries. *Performance Measurement and Metrics*, *10*(1), 49–58.

Noh, Y. (2011). Evaluation of the resource utilization efficiency of university libraries using DEA techniques and a proposal of alternative evaluation variables. *Library Hi Tech*, *29*(4), 697–724.

Pedraja-Chaparro, F., Salinas-Jiménez, J., & Smith, P. C. (2005). Assessing public sector efficiency: Issues and methodologies. *Paper presented at Banca d'Italia Research Department public finance workshop in public expenditure.*

Reichmann, G. (2004). Measuring university library efficiency using data envelopment analysis. *Libri*, *54*(2), 136–146.

Reichmann, G., & Sommersguter-Reichmann, M. (2010). Efficiency measures and productivity indexes in the context of university library benchmarking. *Applied Economics*, *42*(3), 311–323.

Sharma, K. R., Leung, P., & Zane, L. (1999). Performance measurement of Hawaii state public libraries: An application of data envelopment analysis (DEA). *Agricultural and Resource Economics Review*, *28*(2), 190–198.

Shim, W. (2000). Assessing technical efficiency of research libraries. *Advances in Library Administration and Organization*, *17*, 243–339.

Simar, L., & Wilson, P. W. (2000). Statistical inference in nonparametric frontier models: The state of the art. *Journal of Productivity Analysis*, *13*(1), 49–78.

Simar, L., & Wilson, P. W. (2008). Statistical inference in nonparametric frontier models: Recent developments and perspectives. In H. Fried, C. A. K. Lovell, & S. Schmidt (Eds.), *The measurement of productive efficiency and productivity growth* (pp. 421–521). New York, NY: Oxford University Press.

Simon, J., Simon, C., & Arias, A. (2011). Changes in productivity of Spanish university libraries. *Omega*, *39*(5), 578–588.

Smith, P. C. (1997). Model misspecification in data envelopment analysis. *Annals of Operations Research*, *73*, 233–252.

Smith, P. C., & Street, A. (2005). Measuring the efficiency of public services: The limits of analysis. *Journal of the Royal Statistical Society, Series A*, *168*(2), 401–417.

Talluri, S. (2000). A benchmarking method for business-process reengineering and improvement. *International Journal of Flexible Manufacturing Systems*, *12*(4), 291–304.

Thanassoulis, E. (2001). *Introduction to the theory and application of data envelopment analysis: A foundation text with integrated software.* Dordrecht, The Netherlands: Kluwer Academic Publishers.

Troupin, S., Stroobants, J., & Steen, T. (2015). Fiscal consolidation in federal Belgium. *International Review of Administrative Sciences* (forthcoming).

Tulkens, H. (1993). On FDH efficiency analysis: Some methodological issues and applications to retail banking, courts and urban transit. *Journal of Productivity Analysis*, *4*(1/2), 183–210.

van Helden, G. J., & Reichard, C. (2013). A meta-review of public sector performance management research. *Tékhne*, *11*(1), 10–20.
Vitaliano, D. F. (1998). Assessing public library efficiency using data envelopment analysis. *Annals of Public and Cooperative Economics*, *69*(1), 107–122.
Woodbury, K., Dollery, B., & Rao, P. (2003). Is local government efficiency measurement in Australia adequate? An analysis of the evidence. *Public Performance & Management Review*, *27*(2), 77–91.
Worthington, A. (1999). Performance indicators and efficiency measurement in public libraries. *Australian Economic Review*, *32*(1), 31–42.
Worthington, A., & Dollery, B. (2000). Measuring efficiency in local governments' planning and regulatory function. *Public Productivity & Management Review*, *23*(4), 469–485.

Jesse Stroobants received his MS in applied economics (commercial engineer) in 2003, and was a teaching assistant between 2005 and 2010 at the Faculty of Business and Economics of the University of Leuven, Belgium. He currently serves as researcher at the Public Governance Institute at the University of Leuven. His research interests include performance management, performance measurement and efficiency issues in the public sector. He has published, together with Geert Bouckaert, in *Public Organization Review*.

Geert Bouckaert is professor at the Public Governance Institute at the University of Leuven, Belgium. He is currently the President of the International Institute of Administrative Sciences (IIAS) (2013–2016). He was the President of the European Group for Public Administration (EGPA) (2004–2010) and the Chair of the IIAS Programme and Research Advisory Committee (PRAC) (2010–2013). His fields of research and teaching are public management, public sector reforms, performance management and financial management. Among his numerous books and articles is the influential *Public Management Reform. A Comparative Analysis* (Oxford University Press, 2011), co-authored with Christopher Pollitt, now in its 3rd edition. Among other honors, he has received the 2013 ASPA International Public Administration Award of the American Society for Public Administration and holds honorary doctorates at the Université Paul Cézanne Aix-Marseille III, Tallinn Technical University, Corvinus University Budapest, the National School of Political Studies and Public Administration of Romania, University Babes-Bolyai of Romani,a and Université de Lausanne. He is a member of the editorial boards of distinguished journals in his field, for example, *Public Administration Review*.